

DOCTORAL THESIS

Methyl-selective isotope labeling using α -ketoisovalerate
for the yeast *Pichia pastoris* recombinant protein expression system

RIKA SUZUKI

Yokohama City University

Graduate School of Medical Life Science

Department of Medical Life Science

2019

Methyl-selective isotope labeling using α -ketoisovalerate
for the yeast *Pichia pastoris* recombinant protein expression system

RIKA SUZUKI

Yokohama City University

Graduate School of Medical Life Science

Department of Medical Life Science

2019

Member of Judging Committee

Professor (Chair)	Takahisa Ikegami
-------------------	------------------

Associate Professor	Nana Kawasaki
---------------------	---------------

Associate Professor	Mamoru Sato
---------------------	-------------

Visiting Professor	Makoto Arita
--------------------	--------------

Journal of Biomolecular NMR
DOI:10.1007/s10858-018-0192-3
Published online:05/06/2018

Table of contents

Abstract	2
Introduction.....	3
Materials and methods	7
Results and Discussion.....	14
Conclusion.....	19
Figure legends	21
References	28
Supporting Information.....	37

Abstract

Methyl-detected NMR spectroscopy is a useful tool for investigating the structures and interactions of large macromolecules such as membrane proteins. The procedures for preparation of methyl-specific isotopically-labeled proteins were established for the *Escherichia coli* (*E. coli*) expression system, but typically it is not feasible to express eukaryotic proteins using *E. coli*. The *Pichia pastoris* (*P. pastoris*) expression system is the most common yeast expression system, and is known to be superior to the *E. coli* system for the expression of mammalian proteins, including secretory and membrane proteins. However, this system has not yet been optimized for methyl-specific isotope labeling, especially for Val/Leu-methyl specific isotope incorporation. To overcome this difficulty, we explored various culture conditions for the yeast cells to efficiently uptake Val/Leu precursors. Among the searched conditions, we found that the cultivation pH has a critical effect on Val/Leu precursor uptake. At an acidic cultivation pH, the uptake of the Val/Leu precursor was increased, and methyl groups of Val and Leu in the synthesized recombinant protein yielded intense ^1H - ^{13}C correlation signals. Based on these results, we present optimized protocols for the Val/Leu-methyl-selective ^{13}C incorporation by the *P. pastoris* expression system.

Introduction

Methyl groups in proteins are sensitive detection probes for NMR observation of biomolecules because they comprise three equivalent protons, and show characteristically slower transverse relaxation rates than commonly observed amide groups (Tugarinov et al. 2003; Ollerenshaw et al. 2003). The relaxation properties of methyl groups are especially advantageous for analyses of large biomolecules, such as membrane proteins or protein complexes, and various kinds of structural information have been extracted from methyl groups to determine protein structures and interactions or to analyze protein dynamics by NMR (Kay 2011; Kerfah et al. 2015; Wiesner and Sprangers 2015). Another advantage of methyl-observation is that methyl groups can be used as semi-uniform structural probes because methyl-containing amino acids are known to exist at protein cores and intermolecular interaction sites (Bordo et al. 1991; Takahashi et al. 2006). In addition, methyl groups are suitable detection probes of membrane proteins, since they bear more than 50% of methyl-containing amino acids in their transmembrane segments (Liu et al. 2002). Thus, selective isotope labeling of methyl groups is indispensable to analyze methyl groups in large proteins with higher sensitivity and resolution by NMR.

For typical methyl-containing amino acids, Ile, Leu, Val (ILV), a robust method to incorporate ^{13}C into methyl groups and ^2H into non-methyl positions, has been developed using *E. coli* recombinant expression systems, in which metabolic precursors of ILV are supplemented in the culture media to achieve the methyl-selective $^1\text{H}/^{13}\text{C}$ incorporation on a uniform ^2H -labeling background (Goto et al. 1999). The ILV precursors, [4- ^{13}C , 3,3- $^2\text{H}_2$]-labeled α -ketobutyrate (α -KB) for Ile and [4- ^{13}C , 3,4',4',4'- $^2\text{H}_4$]-labeled α -ketoisovalerate (α -KIV) for Val and Leu, are

metabolically converted to the corresponding amino acids retaining ^2H and ^{13}C at their original positions, and these amino acids are eventually used for protein biosynthesis (Fig. 1). Approximately 90% incorporation of ^{13}C into the methyl groups has been achieved by this strategy without severe isotope scrambling into other amino acids. A practical advantage of this method is that these isotopically labeled precursors cost less than the corresponding labeled amino acids.

Although *E. coli* expression systems are widely used, and various isotope labeling strategies have been developed for these expression systems, they rarely produce secretory proteins with multiple disulfide bonds and eukaryotic membrane proteins in functional forms (Monsalve et al. 1999; Bardwell 1994; Leuking et al. 2000; Tate 2001).

Mammalian and insect cells are known to produce these proteins in their active forms, however, yields of the expressed proteins are not necessarily high enough for NMR studies. Another disadvantage for these expression systems is that perdeuteration of the recombinant proteins, which is essential for NMR analyses of large proteins, is so far inapplicable since they cannot survive in fully deuterated medium (Katz and Crespi 1966). In the case of G-protein-coupled receptors, it was reported that a high level of deuteration could be achieved by culturing insect cells with a limited number of ^2H -labeled amino acids (Kofuku et al. 2014).

Yeasts possess attractive features for the expression of secretory and membrane proteins for NMR analyses. A fundamental advantage is that a large number of eukaryotic membrane proteins have been successfully expressed in their active forms in yeast membranes (Reinhart and Kretzler 2006; Lundstrom et al. 2006). Other important features are: (1) vectors for recombinant protein expression are commercially available, and genetic manipulations are as

easy as those of *E. coli* (Rosenfeld 1999; Mattanovich et al. 2012); (2) special equipment is not required for culturing yeasts; and (3) yeast cells can survive in $^2\text{H}_2\text{O}$ and perdeuteration of recombinant proteins can be achieved (Massou et al. 1999; Morgan et al. 2000; Miyazawa-Onami et al. 2013).

The methylotrophic yeast *Pichia pastoris* (*P. pastoris*) is the most widely used yeast to express recombinant proteins (Byrne 2015), and various heterologous proteins could be expressed by *P. pastoris* (Cereghino and Cregg 2000). The commonly used alcohol oxidase 1 (AOX1) promoter is activated by the addition of methanol, and the added methanol is simultaneously consumed by the cells as a sole carbon source in minimal medium, which is typically used for NMR sample preparations.

P. pastoris host strains are classified into three phenotypes (Mut^+ , Mut^S , and Mut^-) with regard to their ability to utilize methanol (Li et al. 2007), and both Mut^+ and Mut^S phenotypes are typically used to express recombinant proteins. The wild-type *P. pastoris* strain, of which the phenotype is Mut^+ , has two alcohol oxidases (AOX1 and AOX2) for methanol consumption, while the Mut^S phenotype, such as the KM71H strain, lacks the AOX1 gene, and its growth on methanol is slower than that of the wild-type. Thus far, a number of difficult proteins, such as albumin (Kobayashi et al. 2000) and membrane proteins including GPCR (Singh et al. 2008; Wetterholm et al. 2008; Hammarberg et al. 2009; Fan et al. 2011; Emami et al. 2013), have been successfully expressed using Mut^S phenotype strains.

Preparations of isotopically labeled proteins for NMR analyses using the *P. pastoris* expression system have been reported for uniform ^{13}C - and ^{15}N -labeling (Laroche et al. 1994; Sugiki et al. 2012), amino acid selective ^{15}N -labeling (Chen et al. 2006), sparse ^{13}C -labeling

using glycerol as a sole carbon source (Liu et al. 2016), and ^2H -labeling (Massou et al. 1999; Morgan et al. 2000). Ile-methyl-selective ^{13}C incorporation on a perdeuterated background using α -KB has been reported so far (Clark et al. 2015; Clark et al. 2017), and the moderate labeling efficiency for δ^1 -methyl groups of Ile residues (45%) was achieved. On the other hand, the efficiency for Val/Leu-methyl-selective ^{13}C -labeling using α -KIV as the precursor was quite low (<5%) and nearly impossible to observe signals from these amino acids (Clark et al. 2015). Recently, Zhang et al. reported that α -KIV is not incorporated into *P. pastoris* cells, and presumed that *P. pastoris* lacks transporters for α -KIV from the extracellular environment to the cytosol (Zhang et al. 2017). They showed that the intensities of methyl resonances originating from Val and Leu decreased by 74% and 72%, respectively, by the addition of 500 mg/L unlabeled Val to culture medium containing 10% (w/w) [U- ^{13}C]-labeled glucose and 90% unlabeled glucose. This result suggested that methyl labeling for Val and Leu residues can be achieved using [^{13}C]-labeled Val, and a similar result was obtained for a different type of yeast, *Kluyveromyces lactis* (*K. lactis*) (Miyazawa-Onami et al. 2013). However, in the case of NMR analyses of large proteins, it is obviously desirable to utilize labeled α -KIV as the precursor for Val/Leu-methyl-selective ^{13}C -labeling on a deuterated background since deuterated Val, such as [4- ^{13}C , 2,3,4',4',4'- $^2\text{H}_5$]-labeled Val, costs a lot.

In this study, to establish a cost-effective method to incorporate ^{13}C into methyl positions of Val and Leu using α -KIV, we explored culturing conditions that allow cellular uptake of α -KIV, and found that acidic pH of the medium is a critical factor for the incorporation of α -KIV into *P. pastoris* cells. By optimizing the culturing method, we successfully incorporated ^{13}C into methyl positions of Val and Leu up to 74% and 10%, respectively. We also established the

culturing method capable of Val-methyl-selective ^{13}C incorporation by the supplementation of unlabeled Leu.

Materials and methods

*Construction of *P. pastoris* cells expressing β -LG*

We used bovine β -lactoglobulin (β -LG) as a model protein in the following study. β -LG is a major component of milk (Sawyer and Kontopidis 2000) and is a secretory protein with a molecular weight of 18K and two intramolecular disulfide bonds. *P. pastoris* is known to express a tremendous amount of this protein (Kim et al. 1997; Denton et al. 1998; Sakurai and Goto 2002), although β -LG can also be expressed by *E. coli* (Ponniah et al. 2010; Loch et al. 2016). β -LG forms a dimer at neutral pH, while it exists as a monomeric form stably at acidic pH (Sakurai and Goto 2007).

The wild-type bovine β -LG gene, corresponding to Y2 through I162, was synthesized (Eurofins Genomics) at the 5' end fused with the Xho1 restriction site and following 18 base pairs (AAAAGAGAGGCTGAAGCT) corresponding to the six C-terminal residues of α -Factor secretion sequence, and at the 3' end fused with a termination codon (TAG) and the Xba1 restriction site. It was inserted into an expression vector for *P. pastoris*, pPICZ α (Invitrogen/Thermo Fisher Scientific), to be fused with an N-terminal α -Factor secretion sequence. The DNA sequence of the ligated vector, pPICZ α - β -LG, was confirmed by an Applied Biosystems genetic analyzer ABI 3130xl (Thermo Fisher Scientific).

The plasmid pPICZ α - β -LG was digested with Pme1, and the linearized vector was transformed

into the cell strains KM71H, following the protocol from the manufacturer (Invitrogen/Thermo Fisher Scientific). The transformants (pPIZ α - β -LG/KM71H) were selected on YPDS (Yeast extract Peptone Dextrose medium with Sorbitol) plates containing 500 μ g/mL Zeocin as an antibiotic, and transformants that produced the maximal amount of β -LG were selected for subsequent large-scale culture.

Optimization of medium pH for α -KIV uptake

Cell culture was started by inoculating 130 mL of BMG (buffered minimal glycerol) medium with four colonies of the transformants. The BMG medium contained 1.34% (w/v) yeast nitrogen base without amino acids, 1% (w/v) ammonium sulfate, 1% (w/v) glycerol, 4×10^{-5} % biotin, and 0.1 M potassium phosphate (pH 6). The culture was shaken at 30 °C and 250 rpm until the turbidity (OD₆₀₀) reached around 15. This starter culture was divided into four culture tubes (~10 mL each) and centrifuged, and the collected cells were resuspended in four types of BM (buffered minimal) media (10 mL each; 1.34% (w/v) yeast nitrogen base without amino acids, 1% (w/v) ammonium sulfate, 4×10^{-5} % biotin, and 0.1 M potassium phosphate), of which pH values were adjusted to either 3.5, 4.0, 5.0, or 6.0. These cultures were shaken for 3 h to allow cells to consume residual glycerol, and a sterile solution of α -KIV (Sigma-Aldrich) was added to each culture to a final concentration of 100 mg/L. After shaking the cultures for 1 h, 0.5 % methanol was added to each culture medium, and the cultures were shaken for 24 h. We sampled 600–800 μ L of each culture medium immediately after the α -KIV addition and at the end of the culture, and the sampled culture media were analyzed by NMR to quantify the α -KIV concentration.

Exploration of factors decreasing recombinant protein expression

Cell culture was started by inoculating 130 mL of BMG medium with four colonies of the transformants. The culture was shaken at 30 °C and 250 rpm until the turbidity (OD₆₀₀) reached around 15. The starter culture was divided into four culture tubes (tube 1 to 4: 10 mL each) and centrifuged. The collected cells were resuspended in 10 mL of two types of media: (1) BM medium at pH 6.0 for tube 1 and 2; (2) BM medium at pH 3.5 for tube 3 and 4. The resuspended cells were cultured for 3 h to allow cells to consume residual glycerol. After this culture period, a sterile solution of α -KIV was added to tube 2 and tube 4 to a final concentration of 100 mg/L. After shaking the culture for another 1 h, 0.5 % methanol was added to each culture, and the cultures were shaken for 24 h. We sampled 600–800 μ L of culture medium from each culture tube immediately after the methanol addition and at the end of the culture, and the sampled culture media were analyzed by NMR to quantify α -KIV and methanol concentrations contained in the media.

Expression of labeled β -LG under an acidic pH condition

Cell culture was started by inoculating 260 mL of BMG medium with four colonies of the transformants. The culture was shaken at 30 °C and 250 rpm until the turbidity (OD₆₀₀) reached around 15. The starter culture was centrifuged at 3,000 \times g for 10 min to collect cells, and the cells were resuspended in the same volume of BM medium, of which pH was adjusted to 3.5. The culture was shaken for 3 h to allow cells to consume residual glycerol, and a sterile solution of [99% 4-¹³C, 98% 3,4',4',4'-²H₄] α -KIV (Cambridge Isotope Laboratory Inc.) was added to the culture to a final concentration of 100 or 200 mg/L. After shaking the culture for 1 h, 0.5 %

methanol was added to the culture, and the culture was shaken for 48 h. To harvest the expressed protein, the culture was centrifuged at 11,899×g, and the supernatant was subjected to the purification steps described below.

Expression of labeled β -LG under a neutral pH condition

Cell culture was started by inoculating 130 mL of BMG medium with four colonies of the transformants. The culture was shaken at 30 °C and 250 rpm until the turbidity (OD₆₀₀) reached 15. The starter culture was centrifuged at 3,000×g for 10 minutes to collect cells, and the cells were resuspended in the same volume of BM medium, of which pH was adjusted to 3.5. The culture was shaken for 3 h to allow cells to consume residual glycerol, and a sterile solution of [4-¹³C, 3,4',4',4'-²H₄] α -KIV was added to the culture to a final concentration of 200 mg/L. The culture was shaken for about 20 h to allow cells to uptake α -KIV, then was centrifuged to collect the cells. The cells were resuspended in the same volume of BMM medium (buffered minimal methanol medium; 1.34% (w/v) yeast nitrogen base without amino acids, 1% (w/v) ammonium sulfate, 4×10⁻⁵% biotin, 0.1 M potassium phosphate, and 0.5% methanol), of which pH was adjusted to 6.0, and the culture was shaken for 24 h to induce the protein expression. The expressed protein was harvested as described below.

Expression of selectively [Val-¹³CH₃]-labeled β -LG

Cell culture was started by inoculating 130 mL of BMG medium with four colonies of the transformants. The culture was shaken at 30 °C and 250 rpm until the turbidity (OD₆₀₀) reached around 15. The starter culture was centrifuged at 3,000 ×g for 10 minutes to collect cells, and the cells were resuspended in the same volume of BM medium of which pH was adjusted to 3.5. The culture was shaken for 3 h to allow cells to consume residual glycerol, and a sterile solution of [4-¹³C, 3,4',4',4'-²H₄] α-KIV was added to the culture to a final concentration of 200 mg/L. The culture was shaken for about 20 h to allow cells to uptake α-KIV and was centrifuged to collect the cells. The cells were resuspended in the same volume of BMM medium adjusted to pH 6.0 and supplemented with 200 mg/L unlabeled Leu. The culture was shaken for 24 h to induce the protein expression. The expressed protein was harvested as described below.

Purification of β-LG

The culture supernatant (~260 mL) was passed through a glass-fiber filter (ADVANTEC, GA-200, 110 nm diameter), and was concentrated to 30 mL by a Minimate TFF System (Pall Corp.) with a 10K molecular weight cut-off membrane. After the concentration, the buffer of the solution was exchanged to an imidazole buffer containing 50 mM imidazole (pH 6.5). The concentrated β-LG was loaded into a column containing 3 mL resin of Q-Sepharose Fast Flow (GE Healthcare) equilibrated with the imidazole buffer. The column was washed with 30 mL of the imidazole buffer, then the elution was allowed by increasing the NaCl concentration stepwise (0.1, 0.2, 0.3, 0.4, and 0.5 M; 9 mL each). β-LG was eluted in the fractions of 0.2 and 0.3 M NaCl, and these eluates were mixed and desalted by dialysis to 2 L of phosphate buffer 1 (20 mM sodium phosphate, pH 2.3) overnight. The desalted β-LG solution was concentrated

by an Amicon Ultra (MWCO 3K; Merck) to 10 mL, and loaded into a Hitrap SP Fast Flow column (1 mL resin volume; GE Healthcare) attached to an AKTA purifier FPLC instrument (GE Healthcare) and equilibrated with phosphate buffer 1. The elution was allowed with a linear gradient of 0 to 1 M NaCl concentration in phosphate buffer 1. β -LG was eluted in the fractions of 0.5 to 0.6 M NaCl, and the fractions were mixed and concentrated to approximately 300 μ L. The buffer of the solution was exchanged to phosphate buffer 2 (20 mM sodium phosphate, pH 2.3, 20 mM NaCl) using an Amicon Ultra.

NMR measurements for culture supernatants

Each sampled culture medium (600–800 μ L) was centrifuged to remove cells. Five hundred μ L of each supernatant, in which 2 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid and 5% $^2\text{H}_2\text{O}$ were added, was subjected to 1D ^1H NMR measurements. The relaxation delay was set to 10 seconds.

NMR measurements for labeled β -LG

The purified β -LG solution, with a concentration of 0.10 to 0.13 mM, was lyophilized and dissolved in the same volume of $^2\text{H}_2\text{O}$. The samples were inserted in 5-mm symmetrical microtubes (Shigemi). NMR experiments were carried out using Bruker Avance III HD 500 and 800 spectrometers equipped with TCI cryoprobes at 25 °C. Spectra were analyzed using the program Topspin 3.1 (Bruker). 2D ^1H - ^{13}C HSQC experiments (Bruker pulse program “heqcetfpgpsi2”) were recorded with a spectral width of 2515 Hz for the carbon dimension (centered at 26 ppm) and 7003 Hz for the proton dimension (centered at 4.7 ppm), and a

recycling delay of 1.2 seconds. Complex data points for direct and indirect dimensions were 1024 and 128, respectively. Val- and Leu-methyl resonances of bovine β -LG were assigned based on the work by Uhrínová et al. (Uhrínová et al. 1998).

MALDI-TOF mass spectrometry for determination of labeling efficiency

Each unlabeled and labeled β -LG sample (typically 0.1 mg) was lyophilized and dissolved in the alkylation reaction buffer containing 50 mM Tris-HCl (pH 8.0) and 6 M Urea. The protein was reduced by 20 mM dithiothreitol at 50 °C for 3 h and was alkylated by the addition of 40 mM iodoacetamide to the solution for 30 min. The buffer of the β -LG solution was exchanged with the digestion reaction buffer containing 50 mM ammonium carbonate (pH 8.0) using a Micro Bio-Spin Chromatography Column (Bio-Rad). The alkylated β -LG was digested by lysyl endopeptidase (Sigma-Aldrich) for 20 h at 37° C. The digestion product was concentrated using a Ziptip pipette tip with C₁₈ resin (Merck), and was eluted stepwise with 5 μ L of 5-50% acetonitrile in 0.1% trifluoroacetic acid. The eluate (1 μ L) was mixed with the same volume of a saturated α -cyano-4-hydroxycinnamic acid solution (Nacalai), and the mixture was loaded on the target plate. MS spectra were obtained with an Autoflex MALDI-TOF-MS (Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight Mass Spectrometry) spectrometer (Bruker). Calibration was performed with Peptide Calibration Standard (Bruker).

Two peptide fragments with amino acid sequences of EAEAYVTQTMK (residues –3-8) and ALPMHIRLSFNPTQLEEQC*HI (residues 142-162; C* corresponds to carbamidemethyl cysteine) were subjected to labeling efficiency determination for Val and Leu, respectively. Incorporation of one [4-¹³C, 3,4',4',4'-²H₄] α -KIV into a Val or Leu residue increases the

molecular mass by 5. The isotope labeling efficiency of the peptides was determined using the following formula:

Labeling efficiency (%)

$$= \{I_{M+5(\text{labeled})} - I_{M+5(\text{unlabeled})}\} / \{I_{M+5(\text{labeled})} - I_{M+5(\text{unlabeled})} + N\} \times 100 \quad (1)$$

where $I_{M+5(\text{labeled})}$ and $I_{M+5(\text{unlabeled})}$ represent the relative peak intensity of the peptide fragment with a molecular mass of M+5 to that of M+0 fragment for the labeled and unlabeled samples, respectively. N represents the number of Val or Leu residues in the fragments.

Results and Discussion

Acidic pH is essential for uptake of α -KIV and incorporation of ^{13}C into methyl groups of Val and Leu residues

It has been shown that *E. coli* cells uptake α -KIV and metabolize it to biosynthesize Val and Leu in standard minimal (M9) medium (Goto et al. 1999), while α -KIV was not converted to Val and Leu by *P. pastoris* cells grown in standard BMG/BMM media typically used for recombinant protein expressions (Clark et al. 2015; Zhang et al. 2017). We first evaluated the α -KIV concentration of the BMG/BMM media during the culturing period using *P. pastoris* KM71H cells by measuring NMR spectra of the culture supernatant. The experiment showed that the α -KIV concentration barely changes during the culture, and suggested that α -KIV in standard BMG/BMM media was not transported into *P. pastoris* cells. To improve the cellular uptake of α -KIV, we modified culture conditions, such as increasing the α -KIV concentration; the addition of dimethyl sulfoxide (Yu and Quinn 1994; André et al. 2006) and PTM1 trace

salts (Isidro et al 2016) to the medium; and using different *P. pastoris* host strains (wild-type (X33) and Mut^S phenotype (KM71H)); however, these factors produced little effect on the α -KIV uptake.

Although Zhang et al. suggested that *P. pastoris* lacks an α -KIV transporter (Zhang et al. 2017), it has been reported that there are several transporters responsible for organic acids and these transporters show pH dependence for the transportation (Jamalzadeh et al. 2012, Souffriau et al. 2012). To evaluate the pH dependence of α -KIV uptake, we cultured *P. pastoris* KM71H cells in α -KIV-containing media at various pH (3.5, 4.0, 5.0, and 6.0). Fig. 2 shows α -KIV uptake by *P. pastoris* cells at different pH conditions. A higher amount of α -KIV was consumed by the cells at acidic pH, suggesting transporters related to the α -KIV uptake were more active at acidic pH than neutral, or a higher amount of the protonated (uncharged) form of α -KIV passed through the plasma membrane at acidic conditions. However, since the acid dissociation constant of α -KIV is lower than 0.9 (Lopalco et al. 2016), the latter mechanism would have a minor contribution.

Based on this pH-dependent uptake of α -KIV, we evaluated the incorporation of isotopically labeled α -KIV into a recombinant protein. β -LG was expressed by KM71H cells that were cultivated in [4-¹³C, 3,4',4',4'-²H₄] α -KIV-containing BMM medium adjusted at two different pH conditions, pH 3.5 and 6.0. Fig. 3 compares the ¹H-¹³C HSQC spectra of β -LG from these cultures. β -LG expressed at pH 6.0 produced only marginal signals, while the protein produced at pH 3.5 provided signals with 10 (Val) / 3.4 (Leu) times higher intensity than those of pH 6.0 on average. This result clearly indicates that the labeled α -KIV was incorporated into the KM71H cells and was metabolically converted to Val and Leu. Unexpected scrambling of the

compound was not found in ^1H - ^{13}C HSQC spectrum of the labeled sample (Fig. S1).

To quantitatively determine the labeling efficiency, we performed MALDI-TOF-MS analysis of the β -LG peptide fragments digested by lysyl endopeptidase. Fig. 4 shows relative signal intensities of β -LG peptide fragments with and without the isotope labeling. The labeling efficiencies of Val and Leu calculated from these mass signals were $57\pm3.5\%$ and $8.4\pm2.1\%$ by a supplementation of 100 mg/L [$4\text{-}^{13}\text{C}$, $3,4',4',4'\text{-}^2\text{H}_4$] α -KIV in the culture, respectively. The labeling efficiencies were further improved to $74\pm2\%$ and $10\pm1\%$ for Val and Leu, respectively, by increasing the labeled α -KIV concentration of the medium to 200 mg/L (Fig. S2).

The labeling efficiency of Val was 7 times higher than that of Leu, which is consistent with the NMR observation. Zhang et al. reported that Leu residues could be labeled by the addition of 500 mg/L labeled Val into the medium containing glucose as a single carbon source (Zhang et al. 2017). In the present case, methanol in the medium may have somehow inhibited the metabolic pathway from Val to Leu.

We also confirmed that this labeling approach was also applicable to a different *P. pastoris* strain, X33 (Mut^+), and the labeling efficiency of Val was similar to that of the KH71H strain. Construction of *P. pastoris* X33 cell lines that express β -LG and the culture method are described in supporting information. The ^1H - ^{13}C HSQC spectrum of [Val/Leu- $^{13}\text{CH}_3$] β -LG expressed by using the X33 strain is shown in Fig. S3 and the labeling efficiencies are described in the figure legend.

The maximum 74% labeling efficiency of Val achieved in this study is substantially higher than that observed in previous trials using *P. pastoris* (Clark et al. 2015), and even higher than the so far highest efficiency of 15% by using the yeast *K. lactis* expression system (Miyazawa-

Onami et al. 2013).

Optimization of the culture condition

An acidic pH culture of *P. pastoris* cells significantly improved the labeling efficiency of Val and Leu using α -KIV, however, this culture condition reduced the yield of the recombinant protein. When α -KIV-containing culture medium of pH 3.5 was used, 5.0 mg of β -LG was obtained from the 1-L culture, which was 50% lower than that from the culture at pH 6.0. Moreover, the culture with twice the amount of α -KIV further decreased the protein yield to 3.6 mg from the 1-L culture. These results indicate that α -KIV somehow affects the recombinant protein expression.

In order to elucidate the reason for the negative correlation between the recombinant protein expression and the incorporation of α -KIV, we analyzed time-dependent change of the methanol concentration in the culture media. Fig. 5 shows the uptake ratios for methanol, which is not only an inducer of protein expression but also a single carbon source during the protein induction phase, calculated from methyl protons intensities of methanol in ^1H -NMR spectra of the culture supernatants. These ratios indicated that methanol uptake was markedly lower in α -KIV-containing culture medium of pH 3.5 than in α -KIV-free culture media and in the culture media of pH 6.0 (Fig. 5). α -KIV uptake apparently inhibited methanol uptake in the culture medium at pH 3.5, and the lower uptake of methanol would have resulted in the low yield of the recombinant protein. Remarkably, however, the acidic pH of culture medium itself (without α -KIV) had little effect on methanol uptake. This result is supported by the facts that the *P. pastoris* cells are known to be viable in acidic conditions (Chiruvolu et al. 1998), and that the

intracellular pH of the yeast *S. cerevisiae* is reported to be maintained around neutral in acidic cultures (Imai and Ohno 1995).

To improve the yield of the Val/Leu-methyl-selective ^{13}C -labeled proteins, we introduced a two-step culture approach, which consists of the initial α -KIV-uptake phase and the subsequent induction phase. The medium for the initial phase contained the labeled α -KIV but no methanol, and the pH of the medium was adjusted to 3.5. At the second phase, the medium was replaced with a fresh methanol-containing medium of pH 6.0 to induce the expression. As a result, approximately 33% of the supplemented α -KIV was incorporated into the cells in 18-20 h of the initial phase. After 20 h of the initial α -KIV-uptake phase, the medium was replaced, and the second induction phase of the culture was performed for 24 h. The yield of the expressed β -LG by this modified culture scheme was approximately 10 mg, and the labeling efficiencies for Val and Leu were calculated as $57\pm0.9\%$ and $14\pm1\%$ by the MALDI-TOF-MS analysis, respectively. This result indicates that the intracellularly incorporated α -KIV in the first step of the culture was pooled in the cells and was eventually used for the recombinant protein production in the second step. A slight (4%) increase in the labeling efficiency of Leu may suggest that a certain amount of the incorporated α -KIV was metabolized into Leu during the first step.

A severe overlap of NMR resonances of large proteins often hampers NMR analyses. To simplify NMR spectra, we attempted to suppress the resonances originating from Leu, by the addition of unlabeled Leu (a final concentration of 200 mg/L) into the culture medium during the induction phase. Fig. 6 shows the Val/Leu-methyl region of the ^1H - ^{13}C HSQC spectrum of β -LG expressed in the presence of unlabeled Leu. Out of 20 expected signals, 18 signals from

the methyl groups of Val were observed. On the other hand, the methyl groups of Leu provided only weak signals that originated from naturally abundant ^{13}C nuclei. The labeling efficiencies of Val and Leu were $56\pm 1\%$ and $<1\%$, respectively, based on the MALDI-TOF-MS analysis. In this study, we supplemented unlabeled Leu to the culture medium for the suppression of Leu resonances. For the application of larger proteins, Val-methyl-selective ^{13}C incorporation on a perdeuterated background will be achieved by the replacement of unlabeled Leu with $[\text{}^2\text{H}_{10}]$ -Leu in deuterated culture.

Conclusion

In this study, we found that the selective incorporation of ^{13}C in methyl groups of Val/Leu residues could be achieved by the *P. pastoris* recombinant protein expression system by lowering the pH of the culture medium to facilitate the cellular uptake of the Val/Leu precursor, α -KIV. The labeling efficiency of Val residues reached 74%, which was much higher than the reported values for *P. pastoris* and *K. lactis* expression systems using the labeled α -KIV (Clark et al. 2015; Miyazawa-Onami et al. 2013), and was close to the value for *E. coli* expression systems (Goto et al. 1999). Although the labeling efficiency of Leu residues was moderate (10%), it is still superior to those using *K. lactis* expression system, even though branched-chain-amino acid-aminotransferase overexpression was conducted (Miyazawa-Onami et al. 2013). We further presented a modified protocol to produce the labeled protein at biologically natural pH conditions (pH 6.0) with twofold higher yield than the original protocol. This protocol was expanded to the protocol for the production of the $[\text{Val-}^{13}\text{CH}_3]$ -labeled protein, which is useful to simplify NMR spectra of larger proteins.

Recently, a number of structures of supramolecular protein complexes were solved by X-ray crystallography as well as cryo-electron microscopy (Shi 2014; Bai et al. 2015; Cheng 2015), and their atomic coordinates are available. On the other hand, dynamical aspects of such supramolecular complexes, which are closely related to their function, can be elucidated using the methyl-TROSY NMR approach (Kay 2011; Kerfah et al. 2015; Wiesner and Sprangers 2015). For this methyl-based NMR analysis, a variety of methyl-selective labeling strategies have been developed so far, such as Met, Ala, and Thr-selective labeling (Fischer et al. 2007; Gelis et al. 2007; Ayala et al. 2009; Isaacson et al. 2007; Velyvis et al. 2012), stereospecific isotope labeling of methyl groups of Val/Leu (Gans et al. 2010), and methyl selective labeling at the Ile- γ 2 position (Ruschak et al. 2010). However, these methyl-selective labeling methods in a perdeuterated background could be applied to only a limited number of protein complexes which can be produced by *E. coli* expression systems. The presented result will widen protein targets which can be analyzed by NMR. The central carbon metabolism and amino acid biosynthesis network of *P. pastoris* have been increasingly investigated recently (Solà et al. 2004; Solà et al. 2007; Förster et al. 2014; Zhang et al. 2017), and this study provides a new insight into the incorporation and metabolism of the amino acid precursor. This kind of knowledge contributes to a better understanding of the carbon metabolism and the amino acid biosynthesis of *P. pastoris*, and provides an alternative and effective route of stable isotope-labeling schemes of difficult eukaryotic proteins, including human membrane proteins.

Figure legends

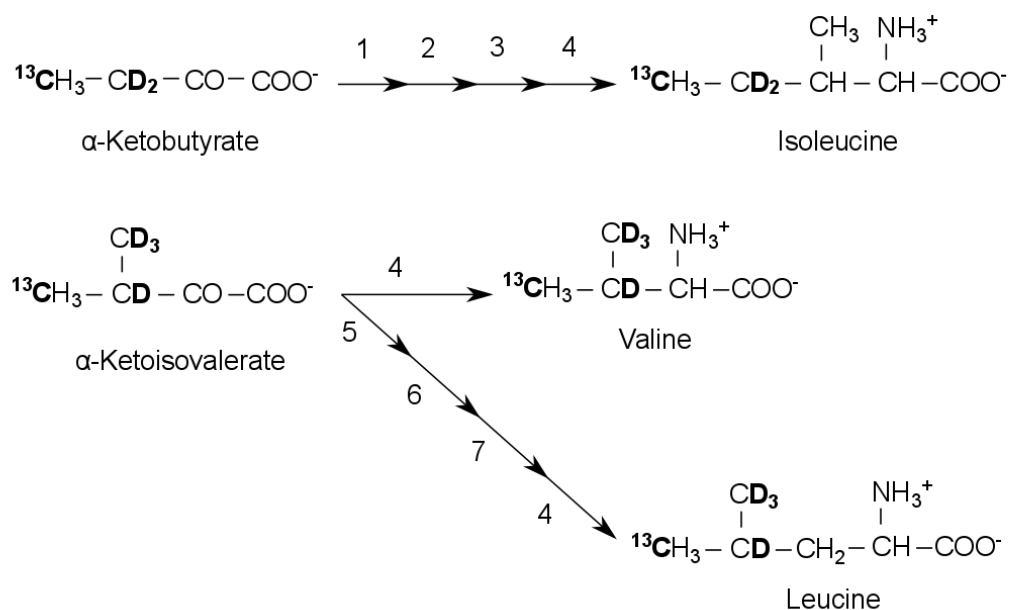


Fig. 1 Simplified biosynthetic pathways of [4- ^{13}C , 3,3- $^2\text{H}_2$]-labeled α -KB and [4- ^{13}C , 3,4',4',4'- $^2\text{H}_4$]-labeled α -KIV. 1: Acetolactate synthase EC 2.2.1.6; 2: Acetohydroxy acid isomerase EC 5.4.99.3; 3: Dihydroxyacid dehydrogenase EC 1.1.1.345; 4: Branched-chain amino acid aminotransferase EC 2.6.1.42; 5: α -isopropylmalate synthase EC 2.3.3.13; 6: Isopropylmalate isomerase EC 4.2.1.33; 7: β -isopropylmalate dehydrogenase EC 1.1.1.85

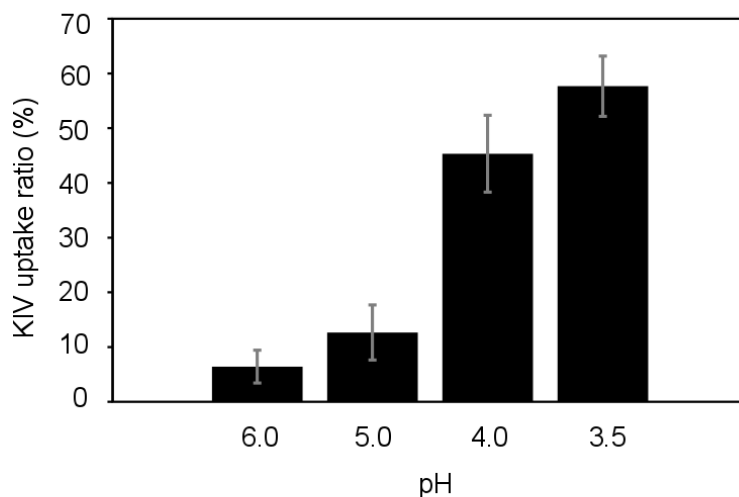


Fig. 2 pH dependence of α -KIV uptake for *P. pastoris* (KM71H strain). Concentration of α -KIV was monitored by NMR measurement of each culture supernatant. The uptake ratio (vertical axis) was calculated as follows :

$$\alpha\text{-KIV uptake ratio (\%)} = (I_{0h}^{KIV} - I_{24h}^{KIV}) / I_{0h}^{KIV} \times 100$$

I_{0h}^{KIV} : The intensity derived from the methyl resonance of α -KIV immediately after the addition of α -KIV to the culture.

I_{24h}^{KIV} : The intensity derived from the methyl resonance of α -KIV after 24 h from the α -KIV addition.

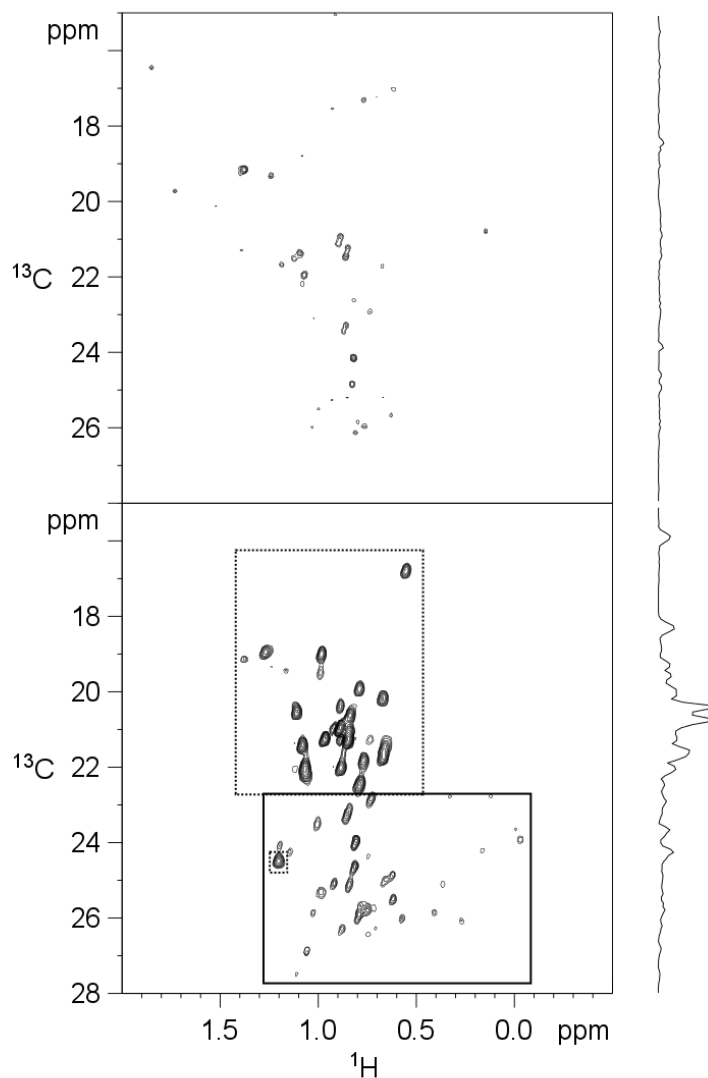


Fig. 3 ^1H - ^{13}C HSQC spectra of [Val/Leu- $^{13}\text{CH}_3$]- β -LG expressed at pH 6.0 (above) and pH 3.5 (below). The projection spectra along the ^{13}C frequency dimension are shown to the right of the 2D spectra. Resonances originating from methyl groups of Val and Leu residues are surrounded by dotted and solid boxes, respectively.

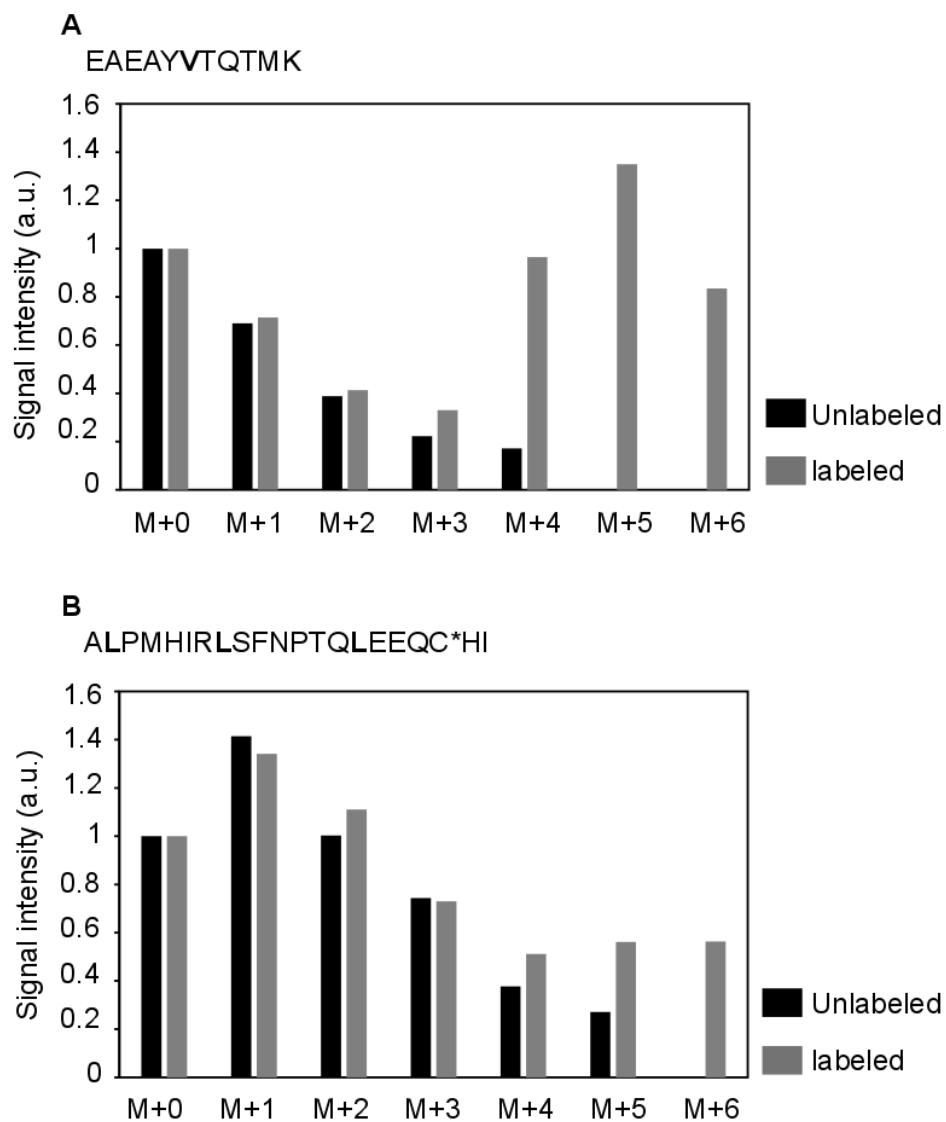


Fig. 4 Relative intensities of the isotope signals from two β -LG peptide fragments (EAEAYVTQTMK: residues -3-8 (A), ALPMHIRLSFNPTQLEEQC*HI: residues 142-162 (B)) observed in MALDI-TOF mass spectra. C* in the latter peptide corresponds to the carbamidemethyl cysteine. Black and gray bars indicate unlabeled β -LG peptide fragments and [Val/Leu- $^{13}\text{CH}_3$]-labeled β -LG peptide fragments (100 mg/L [4- ^{13}C , 3,4',4',4'- $^2\text{H}_4$] α -KIV was

added to the culture), respectively. The Peptide fragment EAEAYVTQTMK has a single Val residue and the peptide fragment ALPMHIRLSFNPTQLEEQC*HI has three Leu residues. Labeling efficiencies of Val and Leu were determined using peptide fragments EAEAYVTQTMK and ALPMHIRLSFNPTQLEEQC*HI, respectively.

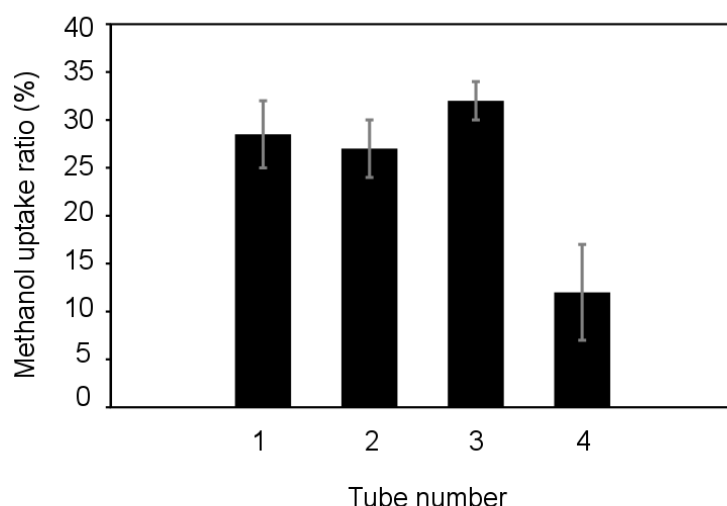


Fig. 5 Methanol uptake of *P. pastoris* (KM71H strain) under different culture conditions (1: BM medium at pH 6.0; 2: BM medium at pH 6.0 with α -KIV; 3: BM medium at pH 3.5; 4: BM medium at pH 3.5 with α -KIV). The methanol concentration was monitored by NMR measurement of each culture supernatant. The uptake ratio (vertical axis) was calculated as follows:

$$\text{Methanol uptake ratio (\%)} = (I_{0h}^{\text{Methanol}} - I_{24h}^{\text{Methanol}}) / I_{0h}^{\text{Methanol}} \times 100$$

I_{0h}^{Methanol} : The intensity derived from the methyl resonance of methanol immediately after the addition of methanol to the culture.

$I_{24h}^{\text{Methanol}}$: The intensity derived from the methyl resonance of methanol after 24 h from the methanol addition.

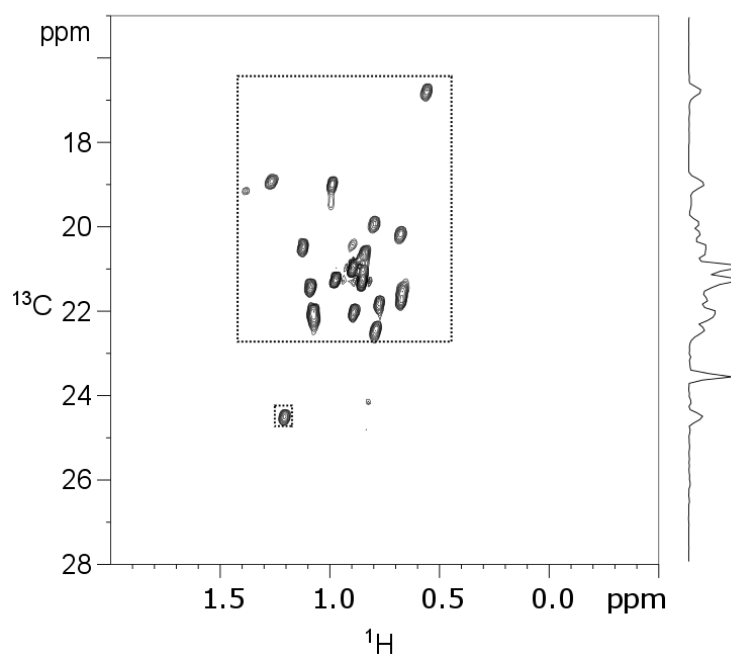


Fig. 6 ^1H - ^{13}C HSQC spectrum of β -LG expressed by the modified labeling scheme with the optional supplementation of unlabeled Leu. The ^{13}C -projection spectrum is shown to the right of the 2D spectrum. Resonances originating from methyl groups of Val residues are surrounded by dotted boxes.

References

- André, N., Cherouati, N., Prual, C., Steffan, T., Zeder-Lutz, G., Magnin, T., Pattus, F., Michel, H., Wagner, R. & Reinhart, C. Enhancing functional production of G protein-coupled receptors in *Pichia pastoris* to levels required for structural studies via a single expression screen. *Protein Sci* **15**, 1115-1126 (2006).
- Ayala, I., Sounier, R., Usé, N., Gans, P. & Boisbouvier, J. An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. *J Biomol NMR* **43**, 111-119 (2009).
- Bai, X. C., McMullan, G. & Scheres, S. H. How cryo-EM is revolutionizing structural biology. *Trends Biochem Sci* **40**, 49-57 (2015).
- Bardwell, J. C. Building bridges: disulphide bond formation in the cell. *Mol Microbiol* **14**, 199-205 (1994).
- Bordo, D. & Argos, P. Suggestions for “safe” residue substitutions in site-directed mutagenesis. *J Mol Biol* **217**, 721-729 (1991).
- Byrne, B. *Pichia pastoris* as an expression host for membrane protein structural biology. *Curr Opin Struct Biol* **32**, 9-17 (2015).
- Cereghino, J. L. & Cregg, J. M. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol Rev* **24**, 45-66 (2000).
- Chen, C. Y., Cheng, C. H., Chen, Y. C., Lee, J. C., Chou, S. H., Huang, W. & Chuang, W. J. Preparation of amino-acid-type selective isotope labeling of protein expressed in *Pichia pastoris*. *Proteins* **62**, 279-287 (2006).
- Cheng, Y. Single-particle cryo-EM at crystallographic resolution. *Cell* **161**, 450-457 (2015).

- Chiruvolu, V., Eskridge, K., Cregg, J. & Meagher, M. Effects of glycerol concentration and pH on growth of recombinant *Pichia pastoris* yeast. *Appl Biochem Biotechnol* **75**, 163-173 (1998).
- Clark, L., Zahm, J. A., Ali, R., Kukula, M., Bian, L., Patrie, S. M., Gardner, K. H., Rosen, M. K. & Rosenbaum, D. M. Methyl labeling and TROSY NMR spectroscopy of proteins expressed in the eukaryote *Pichia pastoris*. *J Biomol NMR* **62**, 239-245 (2015).
- Clark, L. D., Dikiy, I., Chapman, K., Rödström, K. E., Aramini, J., LeVine, M. V., Khelashvili, G., Rasmussen, S. G., Gardner, K. H. & Rosenbaum, D. M. Ligand modulation of sidechain dynamics in a wild-type human GPCR. *Elife* e28505 (2017).
- Emami, S., Fan, Y., Munro, R., Ladizhansky, V. & Brown, L. S. Yeast-expressed human membrane protein aquaporin-1 yields excellent resolution of solid-state MAS NMR spectra. *J Biomol NMR* **55**, 147-155 (2013).
- Denton, H., Smith, M., Husi, H., Uhrin, D., Barlow, P. N., Batt, C. A. & Sawyer, L. Isotopically labeled bovine β -lactoglobulin for NMR studies expressed in *Pichia pastoris*. *Protein Expr Purif* **14**, 97-103 (1998).
- Fan, Y., Shi, L., Ladizhansky, V. & Brown, L. S. Uniform isotope labeling of a eukaryotic seven-transmembrane helical protein in yeast enables high-resolution solid-state NMR studies in the lipid environment. *J Biomol NMR* **49**, 151-161 (2011).
- Fischer, M., Kloiber, K., Häusler, J., Ledolter, K., Konrat, R. & Schmid, W. Synthesis of a ^{13}C -methyl-group-labeled methionine precursor as a useful tool for simplifying protein structural analysis by NMR spectroscopy. *Chembiochem* **8**, 610-612 (2007).
- Förster, J., Halbfeld, C., Zimmermann, M. & Blank, L. M. A blueprint of the amino acid

- biosynthesis network of hemiascomycetes. *FEMS Yeast Res* **14**, 1090-1100 (2014).
- Gans, P., Hamelin, O., Sounier, R., Ayala, I., Durá, M. A., Amero, C. D., Noirclerc-Savoye, M., Franzetti, B., Plevin, M. J. & Boisbouvier, J. Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. *Angew Chem Int Ed* **49**, 1958-1962 (2010).
- Gelis, I., Bonvin, A. M., Keramisanou, D., Koukaki, M., Gouridis, G., Karamanou, S., Economou, A. & Kalodimos, C. G. Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* **131**, 756-769 (2007).
- Goto, N. K., Gardner, K. H., Mueller, G. A., Willis, R. C. & Kay, L. E. A robust and cost-effective method for the production of Val, Leu, Ile(δ 1)methyl-protonated ^{15}N -, ^{13}C -, ^2H -labeled proteins. *J Biomol NMR* **13**, 369-374 (1999).
- Hammarberg, T., Hamberg, M., Wetterholm, A., Hansson, H., Samuelsson, B. & Haeggström, J. Z. Mutation of a critical arginine in microsomal prostaglandin E synthase-1 shifts the isomerase activity to a reductase activity that converts Prostaglandin H₂ into Prostaglandin F_{2 α} . *J Biol Chem* **284**, 301-305 (2009).
- Imai, T. & Ohno, T. The relationship between viability and intracellular pH in the yeast *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **61**, 3604-3608 (1995).
- Isaacson, R. L., Simpson, P. J., Liu, M., Cota, E., Zhang, X., Freemont, P. & Matthews, S. A new labeling method for methyl transverse relaxation-optimized spectroscopy NMR spectra of alanine residues. *J Am Chem Soc* **129**, 15428-15429 (2007).
- Isidro, I. A., Portela, R. M., Clemente, J. J., Cunha, A. E. & Oliveira, R. Hybrid metabolic flux analysis and recombinant protein prediction in *Pichia pastoris* X-33 cultures expressing a

- single-chain antibody fragment. *Bioprocess Biosyst Eng* **39**, 1351-1363 (2016)
- Jamalzadeh, E., Verheijen, P. J., Heijnen, J. J. & van Gulik, W. M. pH-dependent uptake of fumaric acid in *Saccharomyces cerevisiae* under anaerobic conditions. *Appl Environ Microbiol* **78**, 705–716 (2012).
- Katz, J. J. & Crespi, H. L. Deuterated organisms: cultivation and uses. *Science* **151**, 1187-1194 (1966).
- Kay, L. E. Solution NMR spectroscopy of supra-molecular systems, why bother? A methyl-TROSY view. *J Magn Reson* **210**, 159-170 (2011).
- Kerfah, R., Plevin, M. J., Sounier, R. Gans, P. & Boisbouvier, J. Methyl-specific isotopic labeling: a molecular tool box for solution NMR studies of large proteins. *Curr Opin Struct Biol* **32**, 113-122 (2015).
- Kim, T. R., Goto, Y., Hirota, N., Kuwata, K., Denton, H., Wu, S. Y., Sawyer, L. & Batt, C. A. High-level expression of bovine β -lactoglobulin in *Pichia pastoris* and characterization of its physical properties. *Protein Eng* **10**, 1339-1345 (1997).
- Kobayashi, K., Kuwae, S., Ohya, T., Ohda, T., Ohyama, M., Ohi, H., Tomomitsu, K. & Ohmura, T. High-level expression of recombinant human serum albumin from the methylotrophic yeast *Pichia pastoris* with minimal protease production and activation. *J Biosci Bioeng* **89**, 55-61 (2000).
- Kofuku, Y., Ueda, T., Okude, J., Shiraishi, Y., Kondo, K., Mizumura, T., Suzuki, S. & Shimada, I. Functional dynamics of deuterated β_2 -adrenergic receptor in lipid bilayers revealed by NMR spectroscopy. *Angew Chem Int Ed* **53**, 13376-13379 (2014).
- Laroche, Y., Strome, V., De Meutter, J., Messens, J. & Lauwereys, M. High-level secretion and

- very efficient isotopic labeling of tick anticoagulant peptide (TAP) expressed in the methylotrophic yeast, *Pichia pastoris*. *Biotechnology (NY)* **12**, 1119-1124 (1994).
- Leuking, A., Holz, C., Gotthold, C., Lehrach, H. & Cahill, D. A system for dual protein expression in *Pichia pastoris* and *Escherichia coli*. *Protein Expr Purif* **20**, 372-378 (2000).
- Li, P., Anumanthan, A., Gao, X. G., Ilangoan, K., Suzara, V. V., Düzgüneş, N. & Renugopalakrishnan, V. Expression of recombinant proteins in *Pichia pastoris*. *Appl Biochem Biotechnol* **142**, 105-124 (2007).
- Liu, Y., Engelman, D. M. & Gerstein, M. Genomic analysis of membrane protein families: abundance and conserved motifs. *Genome Biol* **3**, research0054 (2002).
- Liu, J., Liu, C., Fan, Y., Munro, R. A., Ladizhansky, V., Brown, L. S. & Wang, S. Sparse ¹³C labelling for solid-state NMR studies of *P. pastoris* expressed eukaryotic seven-transmembrane proteins. *J Biomol NMR* **65**, 7-13 (2016).
- Loch, J. I., Bonarek, P., Tworzydło, M., Polit, A., Hawro, B., Łach, A., Ludwin, E. & Lewiński, K. Engineered β -lactoglobulin produced in *E. coli*: purification, biophysical and structural characterisation. *Mol Biotechnol* **58**, 605-618 (2016).
- Lopalco, A., Douglas, J., Denora, N. & Stella, V. J. Determination of pKa and hydration constants for a series of α -keto-carboxylic acids using nuclear magnetic resonance spectrometry. *J Pharm Sci* **105**, 664-672 (2016).
- Lundstrom, K., Wagner, R., Reinhart, C., Desmyter, A., Cherouati, N., Magnin, T., Zeder-Lutz, G., Courtot, M., Prual, C., André, N., Hassaine, G., Michel, H., Cambillau, C. & Pattus, F. Structural genomics on membrane proteins: comparison of more than 100 GPCRs in 3 expression systems. *J Struct Funct Genomics* **7**, 77-91 (2006).

- Massou, S., Puech, V., Talmont, F., Demange, P., Lindley, N. D., Tropis, M. & Milon, A. Heterologous expression of a deuterated membrane-integrated receptor and partial deuteration in methylotrophic yeasts. *J Biomol NMR* **14**, 231-239 (1999).
- Mattanovich, D., Branduardi, P., Dato, L., Gasser, B., Sauer, M. & Porro, D. Recombinant protein production in yeasts. *Methods Mol Biol* **824**, 329-358 (2012).
- Miyazawa-Onami, M., Takeuchi, K., Takano, T., Sugiki, T., Shimada, I. & Takahashi, H. Perdeuteration and methyl-selective ^1H , ^{13}C -labeling by using a *Kluyveromyces lactis* expression system. *J Biomol NMR* **57**, 297-304 (2013).
- Monsalve, R. I., Lu, G. & King, T. P. Expressions of recombinant venom allergen, antigen 5 of yellowjacket (*Vespula vulgaris*) and paper wasp (*Polistes annularis*), in bacteria or yeast. *Protein Expr Purif* **16**, 410-416 (1999).
- Morgan, W.D., Kragt, A. & Feeney, J. Expression of deuterium-isotope-labelled protein in the yeast *Pichia pastoris* for NMR studies. *J Biomol NMR* **17**, 337-347 (2000).
- Ollerenshaw, J. E., Tugarinov, V. & Kay, L. E. Methyl TROSY: explanation and experimental verification. *Magn Reson Chem* **41**, 843-852 (2003).
- Ponniah, K., Loo, T. S., Edwards, P. J. B., Pascal, S. M., Jameson, G. B. & Norris, G. E. The production of soluble and correctly folded recombinant bovine beta-lactoglobulin variants A and B in *Escherichia coli* for NMR studies. *Protein Expr Purif* **70**, 283-289 (2010).
- Reinhart, C. & Krettler, C. *Structural Genomics on Membrane Proteins*. Chapter **8**. (CRC Press, Boca Raton, 2006).
- Rosenfeld, S. A. Use of *Pichia pastoris* for expression of recombinant proteins. *Methods Enzymol* **306**, 154-169 (1999).

- Ruschak, A. M., Velyvis, A. & Kay, L. E. A simple strategy for ^{13}C , ^1H labeling at the Ile- $\gamma 2$ methyl position in highly deuterated proteins. *J Biomol NMR* **48**, 129-135 (2010).
- Sakurai, K. & Goto, Y. Manipulating monomer-dimer equilibrium of bovine β -lactoglobulin by amino acid substitution. *J Biol Chem* **277**, 25735-25740 (2002).
- Sakurai, K. & Goto, Y. Principal component analysis of the pH-dependent conformational transitions of bovine β -Lactoglobulin monitored by heteronuclear NMR. *Proc Natl Acad Sci USA* **104**, 15346-15351 (2007).
- Sawyer, L. & Kontopidis, G. The core lipocalin, bovine beta-lactoglobulin. *Biochim Biophys Acta* **1482**, 136-148 (2000).
- Shi, Y. A glimpse of structural biology through X-ray crystallography. *Cell* **159**, 995-1014 (2014).
- Singh, S., Gras, A., Fiez-Vandal, C., Ruprecht, J., Rana, R., Martinez, M., Strange, P. G., Wagner, R. & Byrne, B. Large-scale functional expression of WT and truncated human adenosine A_{2A} receptor in *Pichia pastoris* bioreactor cultures. *Microb Cell Fact* **7**, 28 (2008).
- Solà, A., Maaheimo, H., Ylönen, K., Ferrer, P. & Szyperski, T. Amino acid biosynthesis and metabolic flux profiling of *Pichia pastoris*. *Eur J Biochem* **271**, 2462-2470 (2004).
- Solà, A., Jouhten, P., Maaheimo, H., Sánchez-Ferrando, F., Szyperski, T. & Ferrer, P. Metabolic flux profiling of *Pichia pastoris* grown on glycerol/methanol mixtures in chemostat cultures at low and high dilution rates. *Microbiology* **153**, 281-190 (2007).
- Souffriau, B., den Abt, T. & Thevelein, J. M. Evidence for rapid uptake of D-galacturonic acid in the yeast *Saccharomyces cerevisiae* by a channel-type transport system. *FEBS Lett* **586**,

2494–2499 (2012).

Sugiki, T., Ichikawa, O., Miyazawa-Onami, M., Shimada, I. & Takahashi, H. Isotopic labeling of heterologous proteins in the yeast *Pichia pastoris* and *Kluyveromyces lactis*. *Methods Mol Biol* **831**, 19-36 (2012).

Takahashi, H., Miyazawa, M., Ina Y., Fukunishi, Y., Mizukoshi, Y., Nakamura, H. & Shimada, I. Utilization of methyl proton resonances in cross-saturation measurement for determining the interfaces of large protein-protein complexes. *J Biomol NMR* **34**, 167-177 (2006).

Tate, C. G. Overexpression of mammalian integral membrane proteins for structural studies. *FEBS Lett* **504**, 94-98 (2001).

Tugarinov, V., Hwang, P. M., Ollerenshaw, J. & Key, L. E. Cross-correlated relaxation enhanced ^1H – ^{13}C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc* **125**, 10420-10428 (2003).

Uhrínová, S., Uhrín, D., Denton, H., Smith, M., Sawyer, L. & Barlow, P. N. Complete assignment of ^1H , ^{13}C and ^{15}N chemical shifts for bovine β -lactoglobulin: Secondary structure and topology of the native state is retained in a partially unfolded form. *J Biomol NMR* **12**, 89-107 (1998).

Velyvis, A., Ruschak, A. M. & Kay, L. E. An economical method for production of ^2H , $^{13}\text{CH}_3$ -threonine for solution NMR studies of large protein complexes: application to the 670 kDa proteasome. *PLoS One* **7**, e43725 (2012).

Wetterholm, A., Martinez Molina, D., Nordlund, P., Eshaghi, S. & Haeggström, J. Z. High-level expression, purification, and crystallization of recombinant rat leukotriene C₄ synthase from the yeast *Pichia pastoris*. *Protein Expr Purif* **60**, 1-6 (2008).

- Wiesner, S. & Sprangers, R. Methyl groups as NMR probes for biomolecular interactions. *Curr Opin Struct Biol* **35**, 60-67 (2015).
- Yu, Z. W. & Quinn, P. J. Dimethyl sulphoxide: A review of its applications in cell biology. *Biosci Rep* **14**, 259-281 (1994).
- Zhang, M., Yu, X. W., Xu, Y., Jouhten, P., Swapna, G. V. T., Glaser, R. W., Hunt, J. F., Montelione, G. T., Maaheimo, H. & Szyperski, T. ^{13}C metabolic flux profiling of *Pichia pastoris* grown in aerobic batch cultures on glucose revealed high relative anabolic use of TCA cycle and limited incorporation of provided precursors of branched-chain amino acids. *FEBS J* **284**, 3100-3113 (2017).

Supporting Information

Supplemental data includes Material and Methods and Supplemental Figures (S1, S2, and S3) with legends.

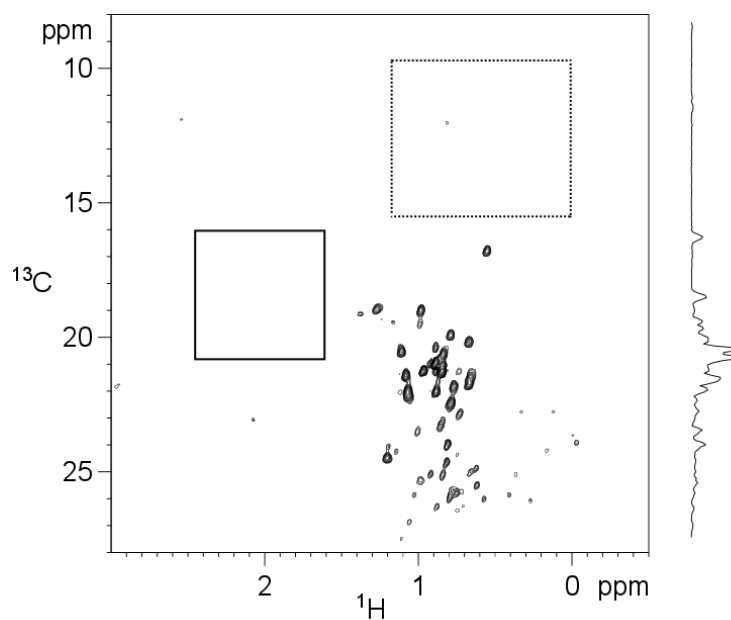
Materials and Methods

*Construction of *P. pastoris* cells (X33 strain) expressing β -LG*

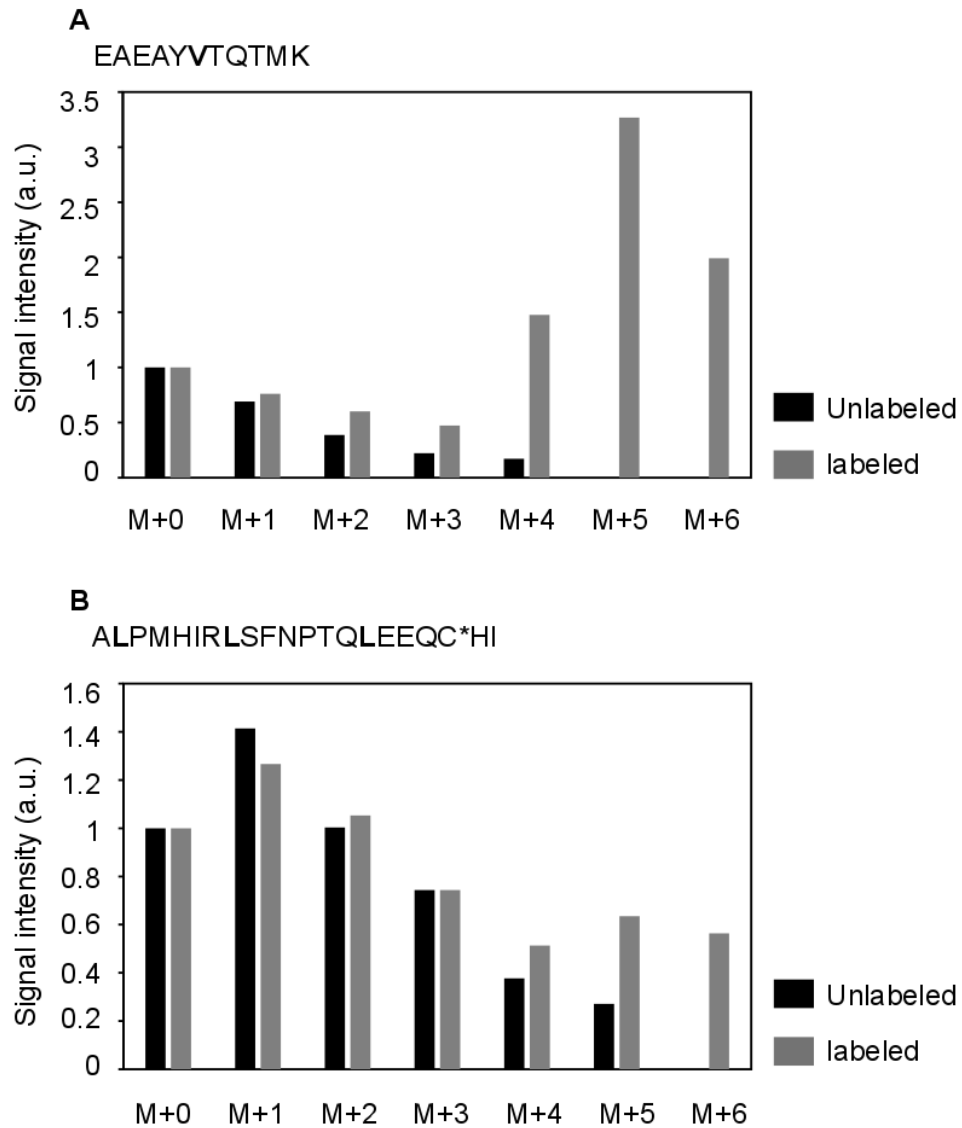
Transformation was carried out in the same manner as for the KM71H strain, and the transformants (pPIZ α - β -LG/X33) were selected on YPDS plates containing 250 μ g/mL Zeocin as an antibiotic. The transformants, which produced the maximal amount of β -LG, were selected for the subsequent large-scale culture.

Expression of labeled β -LG under an acidic pH condition

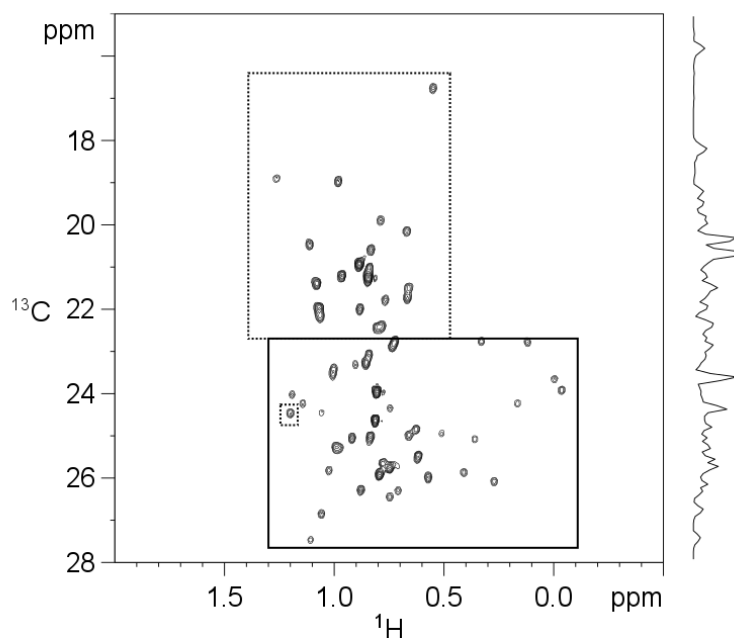
The culturing method was the same as that of the KM71H strain, but the culturing time after the addition of methanol was set to 12 h because the methanol consumption of the X33 strain is faster than that of the KM71H strain.



Supplemental Figure S1. Expanded view of the ^1H - ^{13}C HSQC spectrum of [Val/Leu- $^{13}\text{CH}_3$] β -LG expressed at pH 3.5. The projection spectrum along the ^{13}C frequency dimension is shown to the right of the 2D spectrum. The dotted and solid boxes indicate the spectral regions originating from methyl groups of Ile (δ -methyl) and Met residues, respectively.



Supplemental Figure S2. Relative intensities of the isotope signals from two β -LG peptide fragments (EAEAYVTQTMK: residues -3-8 (A), ALPMHIRLSFNPTQLEEQC*HI: residues 142-162 (B)) observed in MALDI-TOF mass spectra. Black and gray bars indicate unlabeled β -LG peptide fragments and [Val/Leu- $^{13}\text{C}_3$]-labeled β -LG peptide fragments (200 mg/L [4- ^{13}C , 3,4',4',4'- $^2\text{H}_4$] α -KIV was added to the culture), respectively.



Supplemental Figure S3. ^1H - ^{13}C HSQC spectrum of [Val/Leu- $^{13}\text{CH}_3$] β -LG expressed using the X33 strain and medium with 100 mg/L [4- ^{13}C , 3,4',4',4'- $^2\text{H}_4$] α -KIV in the culture. The projection spectrum along the ^{13}C frequency dimension is shown to the right of the 2D spectrum. Resonances originating from methyl groups of Val and Leu residues are surrounded by dotted and solid boxes, respectively.

The labeling efficiencies of Val and Leu calculated from MS analyses were $48\pm 7\%$ and $23\pm 2\%$, respectively, by a supplementation of 100 mg/L labeled α -KIV in the culture. The labeling efficiencies were further improved to $68\pm 4\%$ and $28\pm 3\%$ for Val and Leu, respectively, by increasing the labeled α -KIV concentration of the medium to 200 mg/L.